GPFARM Plant Model Parameters: Complications of Varieties and the Genotype × Environment Interaction in Wheat

G. S. McMaster, J. C. Ascough II, M. J. Shaffer, L. A. Deer–Ascough, P. F. Byrne,

D. C. Nielsen, S. D. Haley, A. A. Andales, G. H. Dunn

Abstract. *The USDA–ARS Great Plains Framework for Agricultural Resource Management (GPFARM) decision support system was developed to assist Great Plains producers in making economically viable and environmentally sound strategic plans for whole farm and ranch systems. A major user requirement for GPFARM is to supply the default plant parameters required to simulate crop growth. Developing this plant parameter database is difficult because varietal differences, caused by a genotype by environment (G × E) interaction, increases parameter uncertainty and variability. This article examines species–based plant parameter sets for simulating winter wheat (*Triticum aestivum *L.) yield responses, explores the significance of the G × E interaction on simulating varietal grain yield, and investigates whether simple adjustments to a species–based plant parameter database can improve simulation of varietal differences across environments. Three plant parameter sets were evaluated against observed yield data for six locations in eastern Colorado: (1) the Default parameter set used best estimates from EPIC–based plant parameter databases, (2) the Dryland Agroecosystems Project (DAP) parameter set further calibrated the default plant parameters against observed yield data for Colorado, and (3) the Theory parameter set modified DAP parameters based on whether irrigated or dryland conditions were simulated. The Theory parameter set simulated yield the best when pooling varieties over environments and locations. However, no parameter set could simulate all the different varietal yield responses to environmental conditions (irrigated or dryland) due to the diverse G × E interactions. The Theory parameter set best simulated the wheat variety TAM 107 across diverse locations, with little bias for either irrigated or dryland conditions. Simple adjustments to a few plant parameters based on whether dryland or irrigated conditions were simulated improved the species–based plant parameter approach used in GPFARM. However, until a better mechanistic representation of the G × E interaction is incorporated into existing plant growth models, opportunities for improving yield response to environmental conditions and management will be limited.*

Keywords. *Crop growth models, Crop yield, Genotype by environment interaction, Plant parameters, Simulation models.*

Agricultural software developers are increasingly delivering products (e.g., decision support systems; simulation models; and budgeting, record keeping, and irrigation/nitrogen/weed control management programs or tools) for use by farmers and ranchers. Unfortunately, most agricultural software is rarely adopted or used on the farm or ranch, especially decision support systems (DSS) and simulation models (Ascough et al., 1999, 2002b). Perhaps the most important reasons for not adopting these products are because they are viewed as too difficult to use and the investment of time and effort to learn, set up, and run the software is not returned in value to the producer.

The USDA–ARS Great Plains Systems Research Unit has developed a decision support system named GPFARM (Great Plains Framework for Agricultural Resource Management). GPFARM 2.5 encompasses stand–alone components such as a user interface, simulation model, and databases (Ascough et al., 2002a; McMaster et al., 2002b; Shaffer et al., 2000) that, when used in conjunction with other components (e.g., farm economic budgeting and multicriteria decision analysis modules), provides a unique decision support tool for farmers and ranchers. GPFARM was designed to provide decision support for use by farmers and ranchers in strategically managing their complex agricultural systems. From the inception of the GPFARM project, producer input on system characteristics and functionality has been solicited. The crop growth component that simulates growth and yield also had to meet these requirements, and still be sufficiently robust to cover the diverse environments, cropping systems, and management practices employed across the Great Plains. Providing a robust crop growth model for producers necessitated that parameters for simulating growth be provided in a

Article was submitted for review in October 2002; approved for publication by the Soil & Water Division of ASAE in May 2003.

The authors are **Gregory S. McMaster,** Agronomist, **James C. Ascough II, ASAE Member Engineer,** Hydraulic Engineer, **Marv J. Shaffer,** Soil Scientist, **Lois A. Deer–Ascough,** Agricultural Engineer, **Allan A. Andales, ASAE Member Engineer,** Visiting Scientist, and **Gale H. Dunn,** Soil Scientist, USDA–ARS Great Plains Systems Research Unit, Fort Collins Colorado; **David C. Nielsen,** Agronomist, USDA–ARS Central Great Plains Research Station, Akron, Colorado; and **Pat F. Byrne,** Associate Professor, and **Scott D. Haley,** Associate Professor, Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado. **Corresponding author:** Dr. Gregory S. McMaster, USDA–ARS, GPSR, 301 S. Howes St., Room 353, Fort Collins, CO 80521; phone: 970–490–8340; fax: 970–490–8310; e–mail: greg.mcmaster@ars.usda.gov.

default database. Unfortunately, correctly parameterizing existing agricultural simulation models and decision support systems has been a major obstacle in successfully applying these products to research and technology transfer (Ahuja and Ma, 2002).

Several approaches have been used to address plant growth parameterization problems. The standard approach collates a set of parameters for species and selected varieties, resulting in a species–based plant parameter set. The user is responsible for parameterizing a particular variety if not included in the data set or if confidence in the default parameters is lacking. Parameter values typically are based on a combination of a single measured value, mean, mid–point of a range, theoretically derived, or even "best–guess or intuition." Difficulties with this approach include the infeasibility of providing values for all possible varieties (e.g., wheat has over 100,000 lines), available measurements are for a limited set of conditions, databases must be continually updated as new varieties are released, and databases are not interchangeable between models because parameters differ.

A rare extension of this approach is to stochastically sample from a distribution of the parameter. This extension is further hampered because even fewer data exist to create a distribution than a single mean value. Efforts to use functional genomics in parameterization are beginning (e.g., Hoogenboom et al., 1997; Welch et al., 2003) but are in their infancy, and much work remains before this approach is useful for software such as GPFARM.

The many causes of variability must be addressed to accurately determine plant parameters. It is well known that plant parameters vary by variety, and that the parameters also vary spatially and temporally across scales. For example, the value of a parameter can vary within a plant and among plants in the "same" environment (e.g., within a small plot) and across plots, locations, and years. To illustrate this, consider a parameter that nearly all crop growth models require: the time required from planting to maturity. Within a wheat plant, different culms reach maturity at different times, and considering only main stems, maturity dates differ within a plot and among plots, locations, and years (McMaster, 1997). Management practices and varying biotic and abiotic factors add further to variability in plant growth parameters. While variance estimates can be associated with each scale of variability, few data exist to characterize the distribution at the species level, much less at the variety level.

Another aspect of variability is the genotype by environment (G × E) interaction that exists at all spatial and temporal scales. Plant breeders and molecular biologists have long recognized the G × E interaction, and indeed this is why yield trials are conducted across locations and years. An essential implication of the G × E interaction is that as the environment changes, genotypes do not respond in the same manner. Yield (or a parameter) for a genotype can respond generally in one of three ways to the environment (Poehlman and Sleper, 1995). One response is for the yield of two genotypes to increase or decrease at a similar rate across an environmental gradient (e.g., nutrient or water availability, temperature). This response shows no G × E interaction. Another response is for the yield of one genotype to increase or decrease more rapidly than another genotype as the environmental gradient changes. This is one type of G × E interaction. The most difficult form of G × E interaction for simulation modeling is known as crossover interaction, in which one cultivar is inferior in yield under one part of the environmental gradient (e.g., low water availability) but superior at another part of the environmental gradient (e.g., high water availability) because yield rankings of varieties differ depending on environmental conditions.

Crop growth models, almost universally, do not explicitly simulate the G × E interaction. Other than pioneering efforts by a few (e.g., Hoogenboom et al., 1997; Welch et al., 2003), we know of no modeling efforts that even attempt to address the G × E interaction issue. Rather, it is assumed that the input value of a parameter for all genotypes has the same response pattern across all environments. Returning to the example of thermal time from sowing to maturity, this assumption ignores the evidence that under different environmental conditions, or management practices altering the environmental conditions, genotypes vary in their relative ranking of thermal time (i.e., there is a G × E interaction; McMaster and Wilhelm, 2003). Considering all varieties, we are likely to find the three G × E responses depending on the particular varieties selected. For simulation modeling purposes, the hope is that for a particular region the normally used "adapted" varieties do not show a significant G × E interaction, and thus avoid the problem.

The crop growth model used in GPFARM (Arnold et al., 1995) is derived from the Water Erosion Prediction Project model (WEPP; Flanagan and Nearing, 1995). A default plant parameter database is provided in which parameters are distinguished only among species, and it is assumed there is no G × E interaction. Adverse ramifications of ignoring the G × E interaction may be diminished as differences in G become smaller (Welch et al., 2003). For example, varieties "adapted" to a region often have greater similarity with each other than with varieties adapted for regions with different environments. Varieties selected for certain management practices (e.g., irrigation or dryland) in a region also often tend to have great similarity in genotype, thus leading to several questions. How accurately does a species–based plant parameter set predict varietal yield in decision support systems such as GPFARM? Are varietal parameters in some form required to accurately simulate yields? Is the G × E interaction a significant problem for a DSS such as GPFARM? Are there simple methods, based on cultural practices (e.g., irrigation or dryland farming), for implicitly incorporating differences in varietal responses to the environment?

The objective of this article is to use different plant parameter sets for winter wheat to: (1) evaluate the adequacy of using a species–based plant parameter database in simulating winter wheat varietal yield responses, (2) explore the significance of the G × E interaction on simulating winter wheat grain yield, and (3) determine whether simple refinements to a species–based database can improve the simulation of winter wheat varietal differences in a DSS such as GPFARM.

GPFARM Crop Growth Component

The GPFARM crop growth component is based on the Water Erosion Prediction Project (WEPP) model crop growth component (Arnold et al., 1995; Deer–Ascough et al., 1998), which is a modified version of the Environmental Policy Integrated Climate (EPIC) crop growth component

Williams et al., 1984, 1989). It has been further modified in GPFARM and incorporates some elements from the Agricultural Land Management Alternatives with Numerical Assessment Criteria (ALMANAC) model (Kiniry et al., 1992). A single model is used for simulating multiple crops by changing model parameters. Stress factors for water and nitrogen are computed using inputs from other independent models within GPFARM.

The crop growth component can be characterized as using the energy– or carbon–driven approach common in plant growth modeling. Potential daily biomass accumulation is based on the interception of light by the canopy (as represented by the LAI and light extinction coefficients) and an energy–to–biomass conversion factor. Limiting abiotic resources are reflected in growth constraint factors (temperature, water, and N) reducing the potential daily biomass accumulation. Carbon and N are partitioned to plant components (e.g., leaves, roots, grain) based on phenological growth stage.

Phenological development of the crop is based on thermal time using daily heat unit accumulation. Daily heat units are computed using the equation:

$$HU_i = \frac{(T_{\max,i} + T_{\min,i})}{2} - T_{base,j} \tag{1}$$

where

HU = heat units (°C) on day i
$T_{\max}$ = maximum temperature (°C) on day i
$T_{\min}$ = minimum temperature (°C) on day i
T_{base} = crop–specific base temperature (°C) of crop j.

No growth occurs at or below T_{base}, and there is no upper temperature limit.

A heat unit index (HUI) ranging from 0 at planting to 1 at physiological maturity is computed as follows:

$$HUI_i = \frac{\sum_{k=1}^{i} HU_k}{PHU_j} \tag{2}$$

where

HUI = heat index for day i
k = counter representing the summation of days
PHU = potential heat units required to reach maturity for crop j.

Several equations are used in determining daily potential biomass production. Interception of photosynthetic active radiation (PAR) is estimated with Beer's law (Monsi and Saeki, 1953):

$$PAR_i = 0.02092(RA_i)(1.0 - e^{-0.65LAI})_i \tag{3}$$

where

PAR = photosynthetic active radiation (MJ m^{-2})
RA = solar radiation (Langleys)
LAI = leaf area index
i = day of the year.

Potential biomass production per day is estimated with the equation (Montieth, 1977):

$$\Delta BP_i = 0.0001(BE_j)(PAR_i) \tag{4}$$

where

ΔBP = potential increase in total biomass on day i (kg m^{-2})
BE = energy to biomass conversion parameter for crop j (kg MJ^{-1}).

Actual daily biomass accumulation is determined by Leibig's Law of the Minimum. The daily potential biomass accumulation (eq. 4) is adjusted daily if one of the plant stress factors (water, N, or temperature) is less than 1.0 using the equation:

$$\Delta B_i = (\Delta BP_i)(REG) \tag{5}$$

where REG is the crop growth regulating factor (the minimum of the water, N, and temperature stress factors) calculated for day i. The adjusted daily total biomass production (ΔB_i) is accumulated through the growing season.

The water stress factor is computed by considering supply and demand in the equation:

$$WS_i = \frac{\sum_{l=1}^{nl} u_{l,i}}{E_{p,i}} \tag{6}$$

where

WS = water stress factor (0–1)
u_l = plant water use in soil layer l (mm)
nl = number of soil layers
E_p = potential plant transpiration (mm)
i = day of the year.

The N stress factor is computed by considering the N demand for biomass production and amount of plant N uptake in the equation:

$$NS_i = \frac{\sum_{l=1}^{nl} V_{l,i}}{N_{p,i}} \tag{7}$$

where

NS = N stress factor (0–1)
V_l = plant N (NO_3–N + NH_4–N) uptake in soil layer l (kg/ha)
N_p = plant N demand
i = day of the year.

N_p is calculated as a percentage of daily total biomass production and varies depending on crop growth stage based on plant parameters BN1, BN2, and BN3 (table 1) for emergence, mid–season, and maturity, respectively.

The temperature stress factor is computed with the equation:

$$TS_i = \sin\left(\frac{\pi}{2}\,\frac{T_{ave,i} - T_{base,j}}{T_{opt,j} - T_{base,j}}\right) \tag{8}$$

where

TS = temperature stress factor (0–1)
T_{ave} = average daily temperature (°C)
T_{opt} = optimum temperature (°C) for crop j
i = day of the year.

Crop yield for annual crops is estimated using the harvest index concept, which is adjusted throughout the growing season according to water stress constraints:

$$YLD_j = (HIA_j)(B_{AG}) \tag{9}$$

where

YLD = crop yield (kg m^{-2}) at harvest for crop j
HIA = adjusted harvest index for crop j

Table 1. List of plant parameters, definitions, and values used in the three plant parameter databases (Default, DAP, and Theory) tested. The Default and DAP parameter sets used the same parameter values for dryland and irrigated conditions. The Theory parameter set used DAP dryland parameter values, and irrigated parameter values different from dryland values are listed. Blank values (—) in the DAP and Theory columns have the same value as to the left within a row.

Variable (Units)	Definition	Default	DAP	Theory
CRIT (°C days)	Growing degree–days from sowing to emergence (annuals)	60	140	
GDDMAX (°C days)	Growing degree–days from sowing to maturity (annuals)	1700	2300	2500
BTEMP (°C)	Base temperature (air) used in calculating growing degree–days	4	0	0
HI (unitless)	Harvest index (dry crop yield / total aboveground biomass)	0.42	0.48	—
XMXLAI (m^2 m^{-2})	Maximum potential leaf area index (LAI)	5	2	4
DLAI (0–1 ratio)	Fraction through growing season when LAI begins to decline	0.8	0.7	0.75
DECFCT (0–1 ratio)	Fraction of canopy remaining after senescence	1	—	—
DROPFC (0–1 ratio)	Fraction of aboveground biomass remaining after senescence	1	—	—
SPRIOD (days)	Period over which senescence occurs	14	—	—
RLAD (unitless)	LAI decline rate parameter	1	—	—
RBIOD (unitless)	Biomass–energy decline rate parameter	10	—	—
PPOP1 (plants m^{-2})	Parameter relating plant population density to maximum LAI	125	—	—
FMLAI1 (0–1 ratio)	Parameter relating maximum LAI for a plant population density	0.6	—	—
PPOP2 (plants m^{-2})	Parameter relating plant population density to maximum LAI	250	—	—
FMLAI2 (0–1 ratio)	Parameter relating maximum LAI for a plant population density	0.95	—	—
BEINP (kg MJ^{-1})	Biomass to energy conversion ratio for a crop	30	—	35
EXTNCT (unitless)	Radiation extinction coefficient	0.65	—	—
OTEMP (°C)	Optimal temperature for plant growth	15	20	—
BB (unitless)	Coefficient relating canopy cover and vegetative biomass	5.2	—	—
BBB (unitless)	Coefficient relating canopy height and vegetative biomass	3	—	—
RDMAX (m)	Maximum rooting depth	1	1.5	1.25
RSR (unitless)	Root biomass to shoot biomass ratio	0.25	—	—
BN1 (0–1 ratio)	Normal fraction of nitrogen in crop biomass at emergence	0.06	—	—
BN2 (0–1 ratio)	Normal fraction of nitrogen in crop biomass at mid–season	0.0231	—	—
BN3 (0–1 ratio)	Normal fraction of nitrogen in crop biomass at maturity	0.0134	—	—
HMAX (m)	Maximum canopy height	0.91	—	—

B_{AG} = cumulative above ground biomass (kg m^{-2}) before senescence occurs.

Harvest index increases nonlinearly from zero at planting using the equation:

$$HI_i = HIO_j\,(HUFH_i - HUFH_{i-1}) \qquad (10)$$

where

HI = harvest index on day i

HIO = harvest index under favorable growing conditions for crop j

$HUFH$ = heat unit factor that affects harvest index for day i and the previous day (i – 1).

The harvest index heat unit is computed with the equation:

$$HUFH_i = \frac{HUI_i}{HUI_i + e^{(6.5-10.0 HUI_i)}} \qquad (11)$$

where $HUFH$ is the harvest index heat unit on day i. The constants in equation 11 are set to allow $HUFH$ to increase from 0.1 at HUI = 0.5 to 0.92 at HUI = 0.9 (Arnold et al., 1995). This is consistent with economic yield development of grain crops, which produce the most economic yield in the second half of the growing season.

Most grain crops are particularly sensitive to water stress near the growth stage of anthesis (Doorenbos and Kassam, 1979). The harvest index is affected by water stress using the following equation:

$$HIA_i = \frac{HI_i}{1.0 + WSYF_j\,(FHU_i)(0.9 - WS_i)} \qquad (12)$$

where

HIA = adjusted harvest index

$WSYF$ = parameter for crop j expressing drought sensitivity (assumed to be a constant 0.01)

FHU = function of crop stage

i = day of the year.

The maximum value for HIA is limited to HI within GPFARM.

Greater detail on the above equations and other equations not discussed (e.g., canopy height, canopy cover, LAI, crop stage factor) can be found in Arnold et al. (1995), Deer–Ascough et al. (1998), and Kiniry et al. (1992).

Methods

Plant Parameter Sets

GPFARM crop–specific plant parameters are kept in a default database to simulate daily growth. Initial default parameters for winter wheat were derived by examining the EPIC, ALMANAC, Soil and Water Assessment Tool (SWAT; Neitsch et al., 2002), and WEPP plant parameter sets that all use the EPIC crop submodel as the foundation for simulating growth. Initial parameter values were determined by taking the midpoint of ranges (if given) or selecting values from regions or varieties closest to the western Central Great Plains where possible. After selecting initial default parameters for each crop species, simulations for typical management practices using generated and historical weather for eastern Colorado were conducted to assess the overall yield response. Critical parameters important in simulating yield (e.g., energy to biomass conversion, harvest index, maximum LAI, thermal time from planting to maturity), found via experience in using the model, were modified to determine if the yield response improved compared to county yield averages. The resulting final default parameter set is denoted as Default in this article; table 1 lists values for winter wheat supplied to GPFARM 2.5 users.

The Default parameter set for winter wheat was further examined and run on a subset of location–treatment–years for the Dryland Agroecosytems Project (DAP) discussed in the Evaluation Data Sets section below (Peterson et al., 2001). Only the wheat–fallow rotation for the summit position at each location was used in calibration (a total of 27 out of 94 location–treatment–years). As in creating the Default plant parameter database, important parameters influencing yield were informally adjusted based on existing data (e.g., McMaster, 1997; McMaster and Smika, 1988; McMaster et al., 1992, 1994) for the variety TAM 107 and "expert opinion" for winter wheat in the western Central Great Plains until simulated yield was improved (Andales et al., 2003). The parameters adjusted were harvest index (HI), maximum potential LAI (XMXLAI), heat unit index when leaf area index begins to decline (DLAI), thermal time from sowing to emergence and maturity (CRIT, GDDMAX), base (BTEMP) and optimum (OTEMP) temperature, and maximum rooting depth (RDMAX). This parameter set is denoted DAP in this article, and values are listed in table 1.

The Theory parameter set (table 1) was developed to address two problems of using species–based plant parameter databases such as the Default and DAP parameter sets. The first problem is that producers often pick varieties they perceive as best adapted for the type of farming they practice, such as irrigated or dryland. Varieties selected for irrigated or dryland conditions can differ greatly in traits, and thus parameter values should likely be changed, but varietal differences are not typically included in plant parameter databases. The second problem relates to simulation of certain processes such as the thermal time from sowing to emergence and maturity. The GPFARM crop growth model assumes a static parameter value (GDDMAX) for the thermal time for the species (or variety), yet it is undeniable that this value should respond to environmental conditions (particularly water stress) other than merely temperature (McMaster, 1997). Therefore, we theorized that simulating grain yield responses might be improved by having different parameters based on dryland and irrigated conditions, which would be an indirect approach to incorporating varietal differences that is a simple refinement easily adapted for all crops. The following parameters were modified for irrigated conditions only: GDDMAX (increased), XMXLAI (increased), DLAI (delayed), BEINP (increased), and RDMAX (decreased). The rationale for direction of change in the parameters is based on fundamental physiological principles. For instance, the onset of leaf senescence (DLAI) and maximum LAI (XMXLAI) is clearly influenced by water availability (McMaster et al., 1992), and this is not accounted for in the model.

Evaluation Data Sets

The primary data used to evaluate the winter wheat yield predictions were based on a two–year study conducted at the Colorado State University Agricultural Research Development and Education Center (denoted ARDEC; 40° 39′ N, 105° 00′ W, 1534 m elevation; fine, smectitic, mesic, Aridic Arguistoll) and the USDA–ARS Central Great Plains Research Station in Akron, Colorado (denoted Akron; 40° 09′ N, 103° 09′ W, 1384 m elevation; fine, smectitic, mesic, Pachic Arguistoll), both initiated in the fall of 1999. Twelve winter wheat varieties (or 10 varieties for year 1 at Akron; listed in fig. 1) differing in presumed heat and drought tolerance were grown under dryland and irrigated conditions. Most varieties are commonly used in this region, but several are adapted to other environments (e.g., Norstar and Siouxland). The experimental design was a split plot with dryland/irrigated conditions the main plot factor and variety the subplot. Replications differed with locations and year: two replications for 1999–2000 at both locations, and three (Akron) and four (ARDEC) replications for 2000–2001. Nitrogen fertilizer was applied at planting to meet recommended levels based on soil tests prior to planting.

Other data were obtained from the long–term DAP study (Peterson et al., 2001) for three locations in eastern Colorado (near Sterling, Stratton, and Walsh) initiated in 1986. Wheat–fallow, wheat–corn–fallow, and wheat–corn–millet–fallow rotations were grown under no–tillage management at different topographic positions of a catena. The winter wheat variety TAM 107 was used through 1998. Nitrogen and phosphorus fertilizer was applied at planting to meet recommended levels based on soil tests prior to planting.

The final data set was from a 6–year study conducted at the Colorado State University Horticultural Farm in Fort Collins, Colorado (denoted Hortfarm; 40° 36′ N, 104° 59′ W, 1515 m elevation) initiated in 1992 on a Nunn clay loam soil (fine, smectitic, mesic, Aridic Arguistoll; McMaster et al., 2002a). A split–plot design, with tillage being the main plot and residue rate being the subplot, with four replications was used with the cultivar TAM 107 in a wheat–fallow cropping system. Soil tests prior to planting indicated that only in the last year was N and P fertilizer (38 kg N ha^{-1} as 32% urea ammonium nitrate solution, and 9.5 kg P ha^{-1} as liquid ammonium polyphosphate: 10–34–0) required to meet recommended levels.

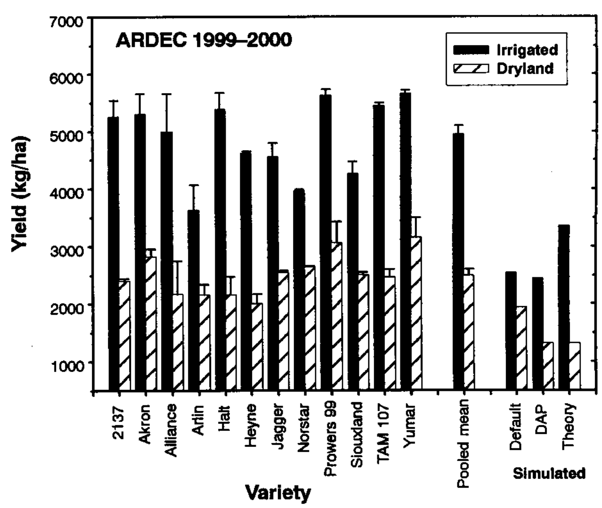

Figure 1. Observed grain yield (with 1 SE bar) for 1999–2000 at ARDEC, Colorado, for dryland and irrigated treatments compared to simulated grain yield for different parameter data sets. Simulated values are the three rightmost sets of bars.

STATISTICAL ANALYSIS

Three statistics were used to evaluate simulation results: (1) paired t–test to determine if differences existed between observed and simulated yield, (2) percentage of simulated yields within 20% of the observed yield (based on producer expectations of GPFARM yield predictions within about 20% of observed), and (3) root mean square error (RMSE), with associated sum of the residuals (SRES) and sum of the absolute residuals (SARES) as described by McMaster et al. (1992). The SRES and SARES measures give an indication of the variability around the mean and tendency for prediction bias and are calculated as:

$$\text{SRES}=\sum_{i=1}^{n}(OBS_i - SIM_i) \qquad (13)$$

$$\text{SARES}=\sum_{i=1}^{n}|OBS_i - SIM_i| \qquad (14)$$

where OBS_i and SIM_i are the observed and simulated yield, respectively. If SRES is small compared to SARES, then errors tend to cancel. If SRES and SARES are large, then the model tends to overpredict yield if SRES is negative and underpredict yield if SRES is positive.

RESULTS AND DISCUSSION

Varietal grain yield responses to different environments, both individually and pooled together, were examined by using the ARDEC and Akron data sets (figs. 1–4). Yield varied with years and two different environmental conditions (dryland and irrigated). As expected, irrigated yields were greater than dryland yields by an average of 56% over the two years and locations.

The ability of species–based data sets (i.e., Default, DAP, and Theory) to simulate yield responses across environments can first be assessed against the observed pooled means (figs. 1–4). When considering the Default and DAP parameter sets derived primarily for dryland conditions, the observed pooled mean clearly was underpredicted in the irrigated and dryland treatments at ARDEC, and the irrigated treatment at Akron was slightly underpredicted (table 2). The Theory parameter set, by introducing different parameters depending on dryland or irrigated conditions, resulted in better simulations of irrigated treatments than either the Default or DAP parameter sets. This suggests that the parameters used in the relationships leading to yield prediction vary depending on the environment.

Pooling the varieties gives an indication of how varieties generally respond to different environments, but can obscure differences based on the particular varieties examined. Using species–based parameter sets (i.e., Default, DAP, and Theory) to simulate varietal differences assumes that merely changing parameters for a variety will work equally well across environments. In other words, the assumption is that there is no G × E interaction and only changing the

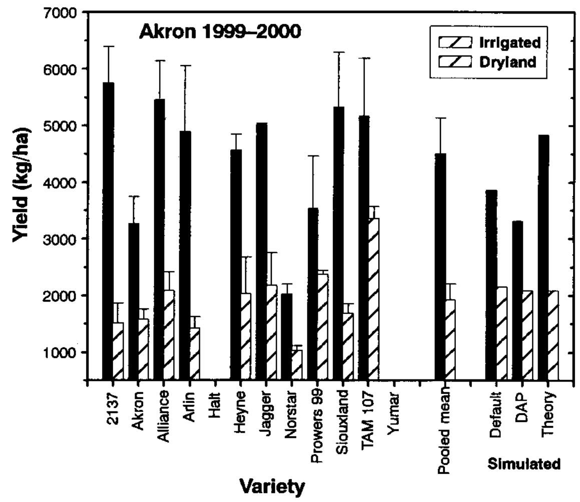

Figure 3. Observed grain yield (with 1 SE bar) for 1999–2000 at Akron, Colorado, for dryland and irrigated treatments compared to simulated grain yield for different parameter data sets. Simulated values are the three rightmost sets of bars. Varieties Halt and Yumar were not grown in 2000.

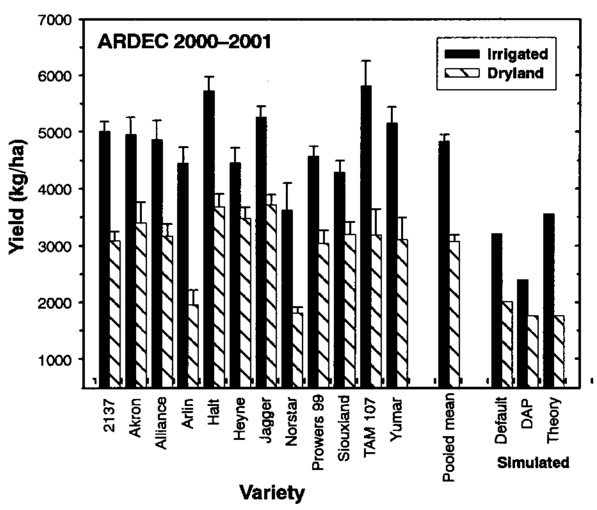

Figure 2. Observed grain yield (with 1 SE bar) for 2000–2001 at ARDEC, Colorado, for dryland and irrigated treatments compared to simulated grain yield for different parameter data sets. Simulated values are the three rightmost sets of bars.

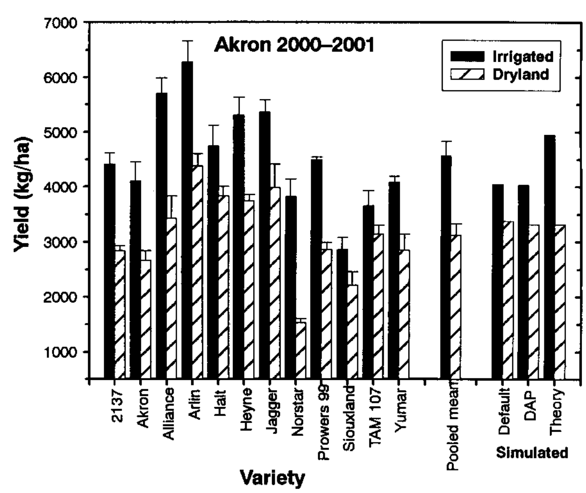

Figure 4. Observed grain yield (with 1 SE bar) for 2000–2001 at Akron, Colorado, for dryland and irrigated treatments compared to simulated grain yield for different parameter data sets. Simulated values are the three rightmost sets of bars.

Table 2. Statistical results of simulating grain yield (kg/ha) at ARDEC and Akron locations using the three plant parameter data sets (table 1). Observed yield was calculated using the mean of 12 cultivars for a site–year treatment. Simulation results are presented in figures 1–4.

Statistical Measure[a]	Default	DAP	Theory
Irrigated and dryland treatments			
RMSE	1158	1464	2761
SRES	6361	8821	4313
SARES	7297	9523	6430
Paired t–test (P)	0.04	0.02	0.12
Dryland treatments			
RMSE	622	889	889
SRES	1152	2136	2136
SARES	2089	2838	2838
Paired t–test (P)	0.43	0.28	0.28
Irrigated treatments			
RMSE	1514	1870	1056
SRES	5208	6685	2177
SARES	5208	6685	3592
Paired t–test (P)	0.06	0.04	0.37

[a] RMSE = root mean square error.
SRES = sum of residuals.
SARES = sum of absolute residuals.
P = probability that there is no difference between observed and simulated yield.

parameters is required to capture the varietal yield response. In our case, if the individual varieties are compared to the simulated yield for the three parameter sets, we would expect to see a consistent bias if the particular variety did not match simulated yields for the parameters chosen. This was not found statistically (data not shown), nor is it apparent in figures 1–4.

To further explore the possible importance of a G × E interaction on yield, we examined the percent yield loss [(irrigated – dryland yield)/irrigated yield × 100] rankings to water stress treatments between years at ARDEC and Akron (figs. 5 and 6). Except for two varieties at each location (Norstar and Arlin at ARDEC, Norstar and Prowers 99 for Akron), greater yield loss was observed in 2000 than in 2001. This was expected, as precipitation was less in 2000 than in 2001 (101 mm and 111 mm less precipitation from September through July for ARDEC and Akron, respectively). These results are encouraging for modeling using decision support systems such as GPFARM in that most varieties tended to respond the same, both between years and locations. Norstar was an anomaly at both locations. Perhaps the percent yield loss response of Norstar, which is a cultivar not typically grown in Colorado, may be partly explained as differentially responding to environmental variables other than water because it is adapted/bred for a different region (e.g., photoperiod and vernalization requirements; heat, cold and drought tolerance; etc.). However, both Arlin and Prowers 99 are commonly grown in Colorado, and this cannot explain their different behavior compared to the other varieties. Further, why did Arlin and Prowers 99 behave differently at only one location? Possibly the unexplained slower rate of Arlin seedling emergence in 2001 at ARDEC caused a greater yield loss in 2001, as the importance of seedling emergence is generally recognized (e.g., McMaster et al., 2002a). In 2001 at Akron, rust was unusually severe and snow occurred on 21 May, resulting in about 5 h below freezing, but nothing unusual in response was observed for Prowers 99.

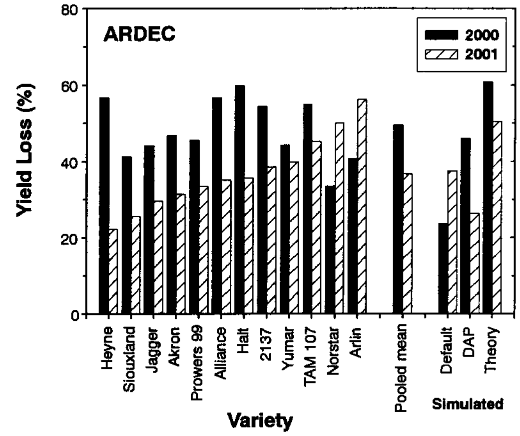

Figure 5. Observed and simulated yield loss for 2000–2001 at ARDEC, Colorado. Simulated values are the three rightmost sets of bars for the three parameters sets evaluated. Percent yield loss = (irrigated – dryland yield)/irrigated yield × 100.

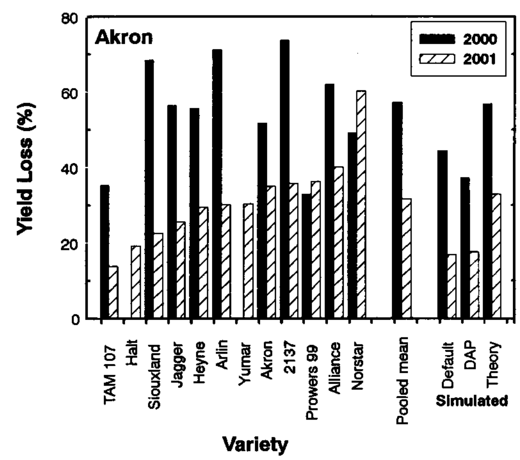

Figure 6. Observed and simulated yield loss for 2000–2001 at Akron, Colorado. Simulated values are the three rightmost sets of bars for the three parameters sets evaluated. Percent yield loss = (irrigated – dryland yield)/irrigated yield × 100. Varieties Halt and Yumar were not grown in 2000.

Regardless of causes, clearly not all varieties responded similarly in yield loss between years or locations, and it is unknown how to explain the observed G × E interaction.

The DAP and Theory parameter sets simulated the pooled mean percent yield loss patterns correctly for both locations (figs. 5 and 6; statistical results not shown). The Default parameter set simulated a very low percent yield loss in 2000 at ARDEC, and therefore the observed pattern between 2000 and 2001 was not correctly simulated. No parameter set was able to correctly simulate both types of varietal yield loss responses observed between years discussed in the preceding paragraph.

The G × E interaction on grain yield can further be seen by examining the change in ranking of percent yield loss of varieties between years (figs. 5 and 6). For instance, at ARDEC, Heyne had among the highest percent yield loss in 2000 (57%), along with Halt (60%), Alliance (57%), TAM 107 (55%), and 2137 (54%), but the lowest percent yield loss in 2001 (22%). Cultivars Norstar and Arlin were the opposite; they had among the lowest percent yield loss in 2000 (34% Norstar and 41% Arlin) and highest in 2001 (50% Norstar and 56% Arlin). At Akron, Siouxland had among the highest percent yield loss in 2000 (68%, along with 71% for Arlin and 74% for 2137), but among the lowest percent yield loss in 2001 (23%). Prowers 99 had a similar change in ranking as Siouxland, but slightly less drastic. Norstar was the opposite (as noted for the ARDEC location above); it had among the lowest percent yield loss in 2000 (49%) and highest in 2001 (60%). This confirms the expected significant G × E interaction in grain yield, but perhaps most importantly, varietal rankings differed between locations, making *a priori* prediction of the G × E interaction even more difficult. Simulating all the various G × E interactions observed in grain yield or percent yield loss was not possible by any parameter set.

The second approach to examining G × E interaction was to plot yield of several varieties across different levels of water stress at each location (figs. 7 and 8). To do this qualitatively, a simplistic assessment of water stress based on growing season precipitation plus irrigation (if applied) was used, and approximately related to a linear scale on the *x*–axis. "Very dry" is represented by dryland treatments in 2000, as precipitation for this year was well below average (109 mm precipitation from September through July at ARDEC compared to a 14–year mean of 295 mm; 285 mm precipitation from September through July at Akron compared to 93–year mean of 367 mm). Since 2001 had more precipitation than 2000 (210 mm precipitation from September through July at ARDEC; 396 mm precipitation from September through July at Akron), dryland treatments for this year were qualitatively characterized as "dry." "Wet" was represented by the mean of the two years of the irrigated treatment. The primary point is that whether comparing varieties within a location or the same variety at different locations, it is clear that varieties differed in their response to water stress, and certain varieties had different response patterns for different locations (caution is needed in comparing locations, because they undoubtedly differed in environmental conditions other than water stress).

For instance, at ARDEC, TAM 107 yield increased approximately linearly in response to water, while Norstar and Arlin yield response to water stress was slightly concave upward, and Heyne was slightly concave downward (fig. 7). At Akron, Norstar responded approximately linearly, TAM 107 was slightly concavely upward, and Arlin and Heyne yield responded concavely downward to water (fig. 8). Although not plotted because the graphs would be unreadable, the other eight varieties showed similarly variable curves. Unless the curves are parallel to each other, there is a G × E interaction. However, despite these differences, it was also apparent that the response of many varieties was somewhat similar so that not all varieties showed a strong G × E interaction. Varieties with significant G × E interaction cannot be simulated using any species–based parameter set that assumes the same response of varieties across a gradient of environmental conditions, as typical of crop simulation models. The next step would be to evaluate the genotypes to ascertain genetic differences related to various G × E responses.

The ARDEC and Akron data showed that not all varieties responded similarly between locations and environments. To further explore the ability of GPFARM to simulate grain yield across locations using different species–based plant parameter sets, one cultivar (TAM 107, one of the most commonly planted varieties in eastern Colorado) was examined. When considering the SRES and SARES values for estimating model bias, the Default parameter set overpredicted dryland yield and underpredicted irrigated yields (table 3 and fig. 9; data not shown for individual analysis of dryland and irrigated yield predictions). The DAP parameter set improved the prediction of dryland yields, although 27 of the 94 treatment–years of the Sterling, Stratton, and Walsh locations were used for calibrating the parameters of the Default parameter set (table 3 and fig. 10). The Theory parameter set improved the irrigated yield predictions (table 3 and fig. 11). Based on the paired t–test and RMSE statistical measures (table 3), the Theory parameter set simulated TAM 107 grain yield across all environments the best, with the DAP parameter set second and the Default parameter set third. Only one measure, the percent of points falling within 20% of the ob–

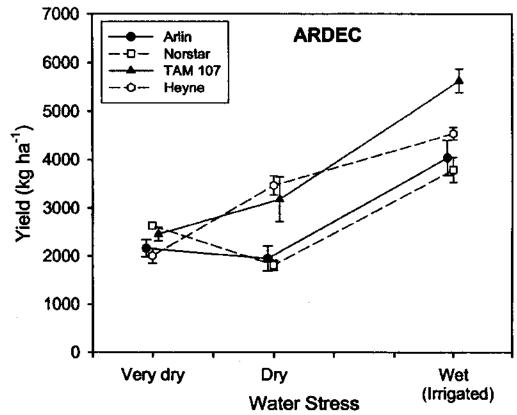

Figure 7. Genotype by environment interaction for different varieties at ARDEC, Colorado. Water stress is based on growing season precipitation from September through July in 1999–2000 ("very dry") and 2000–2001 ("dry"). The "wet (irrigated)" category includes the mean of the two years of irrigated treatments. SE of the mean bars are included for the observed yields.

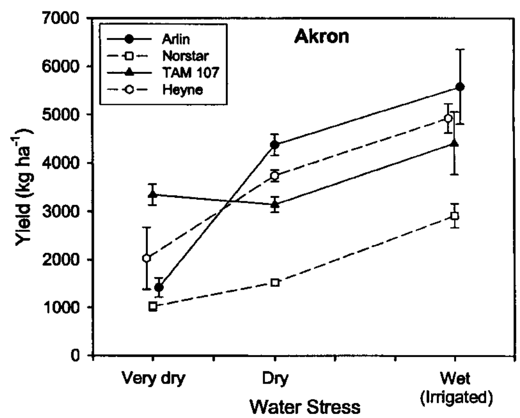

Figure 8. Genotype by environment interaction for different varieties at Akron, Colorada. Water stress is based on growing season precipitation from September through July in 1999–2000 ("very dry") and 2000–2001 ("dry"). The "wet (irrigated)" category includes the mean of the two years of irrigated treatments. SE of the mean bars are included for the observed yields.

Table 3. Statistical results of simulating TAM 107 grain yield (kg/ha) across locations using the three plant parameter data sets (table 1). Simulation results are presented in figures 9–11.

Statistical Measure[a]	Default	DAP	Theory
RMSE	1031	893	811
SRES	43422	9294	4787
SARES	73676	64745	62068
Paired t–test (P)	<0.0001	0.29	0.55
% of points within 20% of observed yield	40.4	36.2	36.2

[a] RMSE = root mean square error.
SRES = sum of residuals.
SARES = sum of absolute residuals.
P = probability that there is no difference between observed and simulated yield.

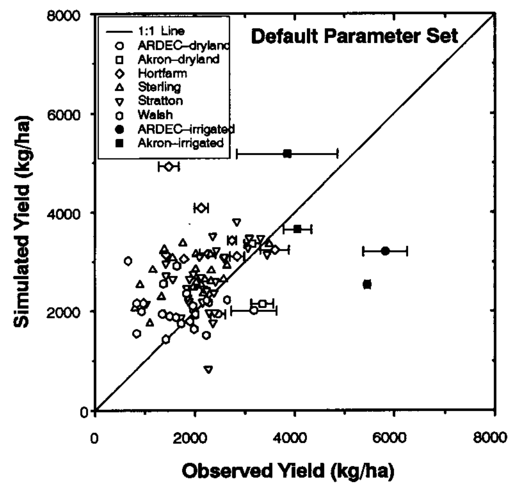

Figure 9. Observed vs. simulated grain yield for cultivar TAM 107 for the ARDEC, Akron, Hortfarm, and DAP (comprised of Sterling, Stratton, and Walsh sites) validation data sets using the Default parameter set. SE bars available only for ARDEC, Akron, and Hortfarm locations. Open symbols are dryland conditions; closed symbols are irrigated conditions.

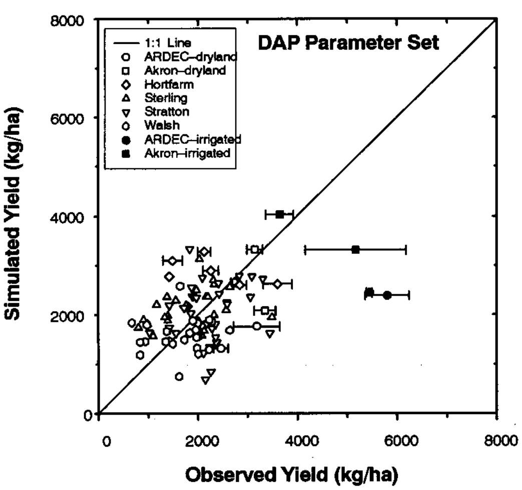

Figure 10. Observed vs. simulated grain yield for cultivar TAM 107 for the ARDEC, Akron, Hortfarm, and DAP (comprised of Sterling, Stratton, and Walsh sites) validation data sets using the DAP parameter set. SE bars available only for ARDEC, Akron, and Hortfarm locations. Open symbols are dryland conditions; closed symbols are irrigated conditions.

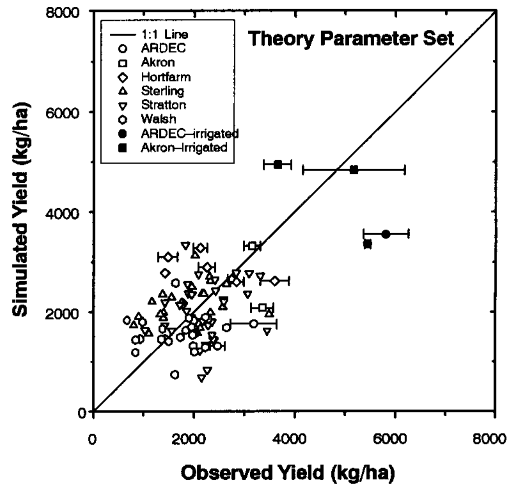

Figure 11. Observed vs. simulated grain yield for cultivar TAM 107 for the ARDEC, Akron, Hortfarm, and DAP (comprised of Sterling, Stratton, and Walsh sites) validation data sets using the Theory parameter set. SE bars available only for ARDEC, Akron, and Hortfarm locations. Open symbols are dryland conditions; closed symbols are irrigated conditions.

served yield, indicated the Default parameter set as being slightly better than the other parameter sets.

TAM 107 was a variety with inconsistent G × E interaction responses to water stress when comparing between locations (figs. 7 and 8). When considering model performance across a range of locations with different management practices, soils, and climates, it appears that a species–based parameter set divided into irrigated and dryland parameter subsets (e.g., Theory parameter set) can simulate TAM 107 grain yield without bias, although whether the scatter around the one–to–one line is acceptable is best left to the individual to decide. Besides the normal causes of simulation error, it is likely that the G × E interaction is significantly contributing to this error. By dividing parameters into two sets for dryland and irrigated conditions, which was easily done *a priori* in our case, a simple alternative was provided for improving plant response relationships with environmental stresses for simulating yield.

Summary and Conclusions

Species–based plant parameter sets cannot reproduce all the complex G × E interactions exhibited by varieties for grain yield. If "unusual" varieties are simulated, the user must be aware that accuracy of results depends greatly on the degree of G × E interaction exhibited by the variety, and certainly significant error is introduced into yield predictions by varietal differences.

Simulating grain yield for one wheat variety (TAM 107), which showed some G × E interaction, across a range of locations and environments was improved by subdividing the parameters into irrigated and dryland values. Undoubtedly, a significant portion of the error was caused by poor mechanistic representation of the G × E interaction in the model. It was not sufficient to merely change parameter values for the potential levels or rates of processes for a variety without also knowing how the level or rate is altered across environmental conditions, which is the implication of the G × E interaction we observed. Subdividing certain parameter values to distinguish environmental conditions (i.e., dryland or irrigated), especially when the model does not adjust the parameter for the environmental conditions, provides a relatively simple improvement for the plant response function to an important environmental condition.

Acknowledgements

The authors thank D. Palic, B. Vandenberg, P. Bartling, D. Edmunds, L. Wiles, S. Canner, N. Ho, D. Hoag, W. M. Frasier, G. Peterson, G. Buchleiter, M. Weltz, L. Ahuja, and J. Hanson for their outstanding efforts in assisting in the development of GPFARM DSS. Partial funding for collecting some field data used in this article was provided by a grant from the Colorado Agricultural Experiment Station.

References

Ahuja, L. R., and L. Ma. 2002. Parameterization of agricultural system models: Current approaches and future needs. In *Agricultural System Models in Field Research and Technology Transfer*, 273–316. L. R. Ahuja, L. Ma, and T. A. Howell, eds. Boca Raton, Fla.: Lewis Publishers.

Andales, A. A, L. R. Ahuja, and G. A. Peterson 2003. Evaluation of GPFARM for dryland cropping systems in eastern Colorado. *Agronomy J.* (in press).

Arnold, J. G., M. A. Weltz, E. E. Alberts, and D. C. Flanagan. 1995. Plant growth component. Chapter 8 in *USDA–Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation*, 8.1–8.41. NSERL Report No. 10. D. C. Flanagan, M. A. Nearing, and J. M. Laflen, eds. West Lafayette, Ind.: USDA–ARS National Soil Erosion Research Laboratory.

Ascough II, J. C., D. L. Hoag, W. M. Frasier, and G. S. McMaster. 1999. Computer use in agriculture: An analysis of Great Plains producers. *Computers and Electronics in Agric.* 23(3): 189–204.

Ascough II, J. C., M. J. Shaffer, D. L. Hoag, G. S. McMaster, L. R. Ahuja, and M. A. Weltz. 2002a. GPFARM: An integrated decision support system for sustainable Great Plains agriculture. In *Sustaining the Global Farm – Local Action for Land Leadership: Selected Papers from the 10th International Soil Conservation Organization (ISCO) Conference*, 951–960. D. E. Stott, R. H. Mohtar, and G. C. Steinhardt, eds. West Lafayette, Ind.: Purdue University and the International Soil Conservation Organization in cooperation with the USDA and Purdue University. See also: http://topsoil.nserl.purdue.edu/nserlweb/isco99/pdf/iscodisc/sustainingtheglobalfarm/p039–ascough.pdf.

Ascough II, J. C., D. L. Hoag, G. S. McMaster, and W. M. Frasier. 2002b. Computer use and satisfaction by Great Plains producers: Implications for software developers. *Agronomy J.* 94:1263–1269.

Deer–Ascough, L. A., G. S. McMaster, J. C. Ascough II, and G. A. Peterson. 1998. Application of generic crop growth model technology to Great Plains conservation systems. ASAE Paper No. 982133. Presented at the ASAE Annual International Meeting. St. Joseph, Mich.: ASAE.

Doorenbos, J., and A. H. Kassam. 1979. Yield response to water. Irrigation and Drainage Paper No. 33. Rome, Italy: United Nations, FAO.

Flanagan, D. C., and M. A. Nearing, eds. 1995. *USDA – Water Erosion Prediction Project Hillslope Profile and Watershed Model Documentation*. NSERL Report No. 10. West Lafayette, Ind.: USDA–ARS National Soil Erosion Research Laboratory.

Hoogenboom, G., J. W. White, J. Acosta–Gallegos, R. G. Gaudiel, J. R. Myers, and M. J. Silbernagel. 1997. Evaluation of a crop simulation model that incorporates gene action. *Agronomy J.* 89: 613–620.

Kiniry, J. R., J. R. Williams, P. W. Gassman, and P. Debaeke. 1992. A general, process–oriented model for two competing plant species. *Trans. ASAE* 35(3): 801–810.

McMaster, G. S. 1997. Phenology, development, and growth of the wheat (*Triticum aestivum* L.) shoot apex: A review. *Advances in Agronomy* 59: 63–118.

McMaster, G. S., and D. E. Smika. 1988. Estimation and evaluation of winter wheat phenology in the central Great Plains. *Agric. and Forest Meteorology* 43(1): 1–18.

McMaster, G. S., and W. W. Wilhelm. 2003. Phenological responses of wheat and barley to water and temperature: Improving simulation models. *J. Agric. Science, Cambridge* 141(1): in press.

McMaster, G. S., J. A. Morgan, and W. W. Wilhelm. 1992. Simulating winter wheat spike development and growth. *Agric. and Forest Meteorology* 60(3–4): 193–220.

McMaster, G. S., W. W. Wilhelm, and P. N. S. Bartling. 1994. Irrigation and culm contribution to yield and yield components of winter wheat. *Agronomy J.* 86: 1123–1127.

McMaster, G. S., D. B. Palic, and G. H. Dunn. 2002a. Soil management alters seedling emergence and subsequent autumn growth and yield in dryland winter wheat–fallow systems in the central Great Plains on a clay loam soil. *Soil and Tillage Research* 65(2): 193–206.

McMaster, G. S., J. C. Ascough II, G. H. Dunn, M. A. Weltz, M. Shaffer, D. Palic, B. Vandenberg, P. Bartling, D. Edmunds, D. Hoag, and L. R. Ahuja. 2002b. Application and testing of GPFARM: A farm and ranch decision support system for evaluating economic and environmental sustainability of agricultural enterprises. *Acta Horticulturae* 593: 171–177.

Monsi, M., and T. Saeki. 1953. Uber den Lictfaktor in den Pflanzengesell – Schaften und sein Bedeutung fur die Stoffproduktion. *Japan J. Botany* 14: 22–52.

Montieth, J. L. 1977. Climate and the efficiency of crop production in Britain. *Phil. Trans. Res. Soc. London B.* 281: 277–329.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, J. R. Williams, and K. W. King. 2002. *Soil and Water Assessment Tool Theoretical Documentation.* Version 2000. TWRI Report TR–191. College Station, Texas: Texas Water Resources Institute.

Peterson, G. A., D. G. Westfall, F. B. Peairs, L. Sherrod, D. Poss, W. Gangloff, K. Larson, D. L. Thompson, L. R. Ahuja, M. D. Koch, and C. B. Walker. 2001. Sustainable dryland agroecosystem management. Technical Bulletin TB01–2. Fort Collins, Colo.: Colorado State University and Agricultural Experiment Station.

Poehlman, J. M., and D. A. Sleper. 1995. *Breeding Field Crops.* 4th ed. Ames, Iowa: Iowa State University Press.

Shaffer, M. J., P. N. S. Bartling, and J. C. Ascough II. 2000. Object–oriented simulation of integrated whole farms: GPFARM framework. *Computers and Electronics in Agric.* 28(1): 29–49.

Welch, S. M., J. L. Roe, and Z. Dong. 2003. A genetic neural network model of flowering time control in *Arabidopsis thaliana*. *Agronomy J.* 95: 71–81.

Williams, J. R., C. A. Jones, and P. T. Dyke. 1984. A modeling approach to determining the relationship between erosion and soil productivity. *Trans. ASAE* 27(1): 129–144.

Williams, J. R., C. A. Jones, J. R. Kiniry, and D. A. Spanel. 1989. The EPIC crop growth model. *Trans. ASAE* 32(2): 497–511.